Early-Warning Mechanism to Detect Signals of Potential Uproar

Ubisoft Challenge

Team 2: Chiara Chianese, Eleonora Di Mario, Giulia Macis, Yeongyeong Jo, Atharva Tidke, Hsin-ni Tsai.

Github repository: https://github.com/Giuliamacis/ubisoft-project

1. INTRODUCTION

In today's world, social controversies—ranging from racism and sexism to cultural appropriation—are more and more scrutinized and condemned. This high awareness poses serious challenges for storytellers, including video game creators, that even with no malicious intent, come up with certain aspects of games that are then perceived as controversial by the audience, sparking the phenomenon known as "gamer uproar."

Even though gamers, who ultimately determine the success or failure of a game through sales, still value mostly the technical aspects of the videogame itself, they can still be influenced by uproars that instead frequently originate from non-gamers, as seen in cases like Assassin's Creed Shadow, which faced backlash for alleged cultural insensitivity and "blackwashing."

For this reason, we agree that addressing this emerging issue is both urgent and complex, as it can significantly help gaming companies safeguard their reputation and brand image.

However, to approach the matter it is important to consider the evolving nature of social dynamics, which makes it harder to build an effective predictive model. To tackle this challenge, we have crafted a feasible, versatile, and data-driven solution.

Our approach proves a reliable early-warning mechanism to identify games at risk of gamer dissatisfaction and uproar, through clear signals detectable by different social platforms.

Our research begins on YouTube, a widely used platform offering a diverse mix of opinions from both gamers and non-gamers. This diversity is critical, as non-gamers are often the source of these controversies. Focusing on the comment sections of video game review videos, we developed and tested two core hypotheses:

H1. Games that are poorly received and criticized generate fewer discussion topics in YouTube comments compared to successful games.

H2. The sentiment of a YouTuber's review influences the overall sentiment of the comment section.

These hypotheses, grounded in psychological theories, form the foundation of our analytical framework. By rigorously testing these hypotheses, we crafted a robust method to detect early signs of potential gamer uproars, providing Ubisoft with actionable insights to safeguard its brand and maintain its audience's trust.

2. THEORETICAL FRAMEWORK

Our study is built on hypotheses grounded in several psychological theories and studies, specifically the broaden-and-build theory of positive emotions, the study of valence asymmetries and the negativity bias theory. These theories offer insights into the effects of positive and negative sentiments on cognitive and social behavior, helping us interpret the observed trends in audience reactions.

2.1 Broaden-and-Build Theory of Positive Emotions

The broaden and build theory describes the form and function of a subset of positive emotions, including joy, interest, contentment, and love. A key proposition is that these positive emotions broaden an individual's range of thoughts and actions. For example: joy stimulates the urge to play, interest stimulates the urge to explore, contentment stimulates the urge to integrate, and love stimulates these actions in a cycle. On the other hand, negative emotions narrow the mindset and lead to specific action tendencies. In the context of our analysis, this principle suggests that when a game or review elicits positive emotions in viewers, they are more likely to engage broadly, discussing a variety of topics and exploring different aspects of the game or review. A second key proposition concerns the consequences of these mindsets: positive emotions and experiences promote discovery of novel and creative actions, ideas and social bonds, which in turn build individual's personal physical, intellectual, social and psychological resources. In the comment sections of a well received game or video, this could manifest as more interactions, richer conversations and a greater tendency to explore different viewpoints. In the context of our study, the broaden-and-build theory supports our hypothesis that positive audience sentiment will correlate with a more dynamic exchange of ideas in the comment section.[1]

2.2 Valence Asymmetries in Processing Positive and Negative Information

The concept of valence asymmetries provides additional insight into the contrasting ways that positive and negative information is processed, examining how people process positive versus negative information differently. The negative information captures attention more readily, is remembered more vividly, and tends to carry more weight in evaluations, while positive information is processed more swiftly and generates broader associative thinking. According to the theory, negative information has a stronger impact on attention. Individuals often take longer to process negative traits or events, indicating a stronger focus on unfavorable aspects. From the point of view of memory, negative information is typically remembered in greater detail. Also, in the perception sphere negative impressions act in a strong and persistent way, in fact negative impressions tend to have a stronger influence on overall judgments about people. Instead, positive information is generally processed more quickly than negative information, allowing a faster response. Another important point, especially for our analysis, is that positive information tends to activate and stimulate associations, making it easier to connect with other concepts and recognizing congruent information. In relation to our analysis, the valence asymmetry model

[1] **Fredrickson, B. L. (2004).** The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359*(1449), 1367–1377. https://doi.org/10.1098/rstb.2004.1512

supports the idea that positive and negative audience reactions in the comment sections of a video will differ in structure and focus. When a review or game generates negative sentiment, the results will be a narrower discussion. On the other hand, positive sentiment encourages faster and broader associations, leading to a more diverse range of topics focusing less on one element, as anticipated by our hypothesis.[2]

2.3 Negativity Bias Theory

The negativity bias theory further substantiates our analysis by explaining why negative sentiments often dominate over positive ones. The negativity bias, also known as the negativity effect, is a cognitive bias that asserts that even when positive or neutral things of equal intensity occur, things of a more negative nature have a greater effect on one's psychological state and processes than neutral or positive things. In other words, something very positive will generally have less of an impact on a person's behavior and cognition than something equally emotional but negative. This theory identifies four specific manifestations of negativity bias. The first one is the negative potency. The principle of negative potency asserts that, given negative and positive events of equal objective magnitude, the negative event is more potent and of higher salience than its positive counterpart. In our study, this principle explains why it's more likely that negative aspects of a video game or video could reach virality and cause backlash. The second one is the greater steepness of negative gradients. Negative events are perceived as increasingly more negative than positive events, the closer one gets, spatially or temporally, to the event itself. In other words, the perception of negativity intensifies more rapidly when approaching negative events, while positive emotions tend to increase more gradually. There is a steeper negative gradient than positive gradient. Also this principle gives us a valid explanation about the virality of bad content and it justifies the rapidity with which backlash is generated. Negativity dominance is the third principle of the theory. It describes the tendency for the combination of positive and negative events or items to incline towards an overall more negative interpretation, in other words the whole is more negative than the sum of its parts. Also, this principle also provides a theoretical foundation for our study, illustrating why, once a discussion in the comment section of a video turns negative, it tends to stay negative or continues to dominate over positive discourse. The fourth and last one is the negative differentiation according to which negative information tends to be represented in more complex and varied cognitive forms than positive information. For instance, research indicates that negative vocabulary is more richly descriptive than that of positive vocabulary. It's also thanks to this principle that we were able to recognize negative topics and collect data about them. Furthermore, there appear to be more terms employed to indicate negative emotions than positive emotions. In the framework of

[2] **Unkelbach, C., Alves, H., & Koch, A. (2020).** Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. *Advances in Experimental Social Psychology, 62*, 115-180. https://doi.org/10.1016/bs.aesp.2020.04.005

our study, the negativity bias theory supports our hypothesis that when games or reviews are perceived negatively, comments will focus primarily on critical or negative aspects, eclipsing potential positive discussions, and making the scope of the conversationn narrower.[3]

2.4 Social Proof Theory

It is considered a psychological and social phenomenon wherein people copy the actions of others in choosing how to behave in each situation. The term was coined by Robert Cialdini in his 1984 book Influence: Science and Practice.[4]

This phenomenon is particularly powerful when the source of influence is perceived as credible, knowledgeable, or similar to the individual. The theory explains this through the mechanisms of credibility, expertise, and similarity:

- Credible sources are trusted because they are perceived to have accurate and reliable information, making their opinions more persuasive.
- Expertise adds to this trust, as individuals believe that knowledgeable figures are better equipped to make accurate judgments.
- Similarity further strengthens the effect because people are more likely to align with the opinions of those they see as sharing their values, experiences, or preferences. These elements together create a strong foundation for social proof, as individuals defer to those they view as credible, expert, or relatable to reduce uncertainty and make decisions more confidently.

This theory allows us to underpin theoretically the hypothesis that we formulated: the sentiment of a YouTuber's review influences the overall sentiment of the comment section.

Youtubers are usually considered as trusted authorities in the gaming domain, they usually showcase their knowledge on the matter, making them credible and with expertise. The strong parasocial relationships viewers form with YouTubers amplify this effect, as followers feel a personal connection and trust in their opinions. Consequently, the sentiment expressed by a well-known YouTuber cascades through their audience, influencing the overall sentiment in the comment section and, ultimately, public perception of the game. This behavior, rooted in social proof, underscores the critical role of YouTubers in shaping gaming narratives and market success.

In summary, these theories provide a robust foundation for our hypothesis. The broaden-and-build theory explains why positive sentiment is associated with a wider range of discussion topics, as positive emotions promote cognitive expansion and curiosity. The valence asymmetry concept highlights how positive and negative information

[3] **Rozin, P., & Royzman, E. B. (2001).** Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2

[4] Cialdini, Robert B. *Influence: Science and Practice*. HarperCollins, 1984

processing differences contribute to varied engagement styles in comment sections, while negativity bias theory explains why negative sentiment leads to more focused, narrow discussions dominated by those negative aspects. Together, these theories provide a strong theoretical support to the results we expect from the analysis. [5]

3. DATA ANALYSIS:

To test and verify our hypotheses, we started by choosing 20 different video games: 10 well-received and 10 poorly received. In order to make a reasonable selection of the video games, we referred to two online platforms: Metacritic and OpenCritic. These platforms are used by players to leave reviews about games and based on that, games are scored. Metacritic was selected because it is perceived by the audience as a trusted platform and thanks to that it has extensive coverage. On the platform both professional critics and users can leave a review about a game. OpenCritic was chosen for its focus specifically on video games. It takes reviews from a wide range of individual critics on various gaming platforms (like PlayStation, Xbox, PC) giving a balanced view across platforms. By combining data from both Metacritic and OpenCritic, we gain a balanced understanding of each game.

After gathering scores, we finalized our game selections in the "poorly received" category with titles like:

The Last of Us Part II (user score: 5.8), *Atomic Heart* (user score: 6.9, critic score: 74), *Suicide Squad: Kill the Justice League* (user score: 3.4, critic score: 59), *Call of Duty: Modern Warfare III* (user score: 2.2, critic score: 58), *Battlefield 2042* (user score: 2.3, critic score: 65), *Concord* (user score: 1.7, critic score: 64), *Fallout 76* (user score: 2.9, critic score: 54), *Kingdom Come: Deliverance* (critic score: 70), *Balan Wonderworld* (user score: 5.2, critic score: 48) and Assassin’s creed shadows.

For the "well received" category:

Space Marine 2 (critic score: 81), *GTA 5* (user score: 8.5,, critic score: 92), *Read Dead Redemption 2* (user score: 8.9, critic score:96), *Players Unknown Battlegrounds* (critic score: 77), *Legend of Zelda tears of the kingdom* (user score: 8.8,, critic score: 96), *Super Mario Bros wonder* (user score: 9, critic score: 91), *Alan wake 2* (user score: 8.5, critic score: 89), *Apex Legends* (user score: 6.6, critic score: 81), *Crisis Core: final fantasy 7* (user score: 7.9, critic score: 79) and *Uncharted 4: a thief’s end* (user score: 8.8, critic score: 93).[6][7]

For each game, we chose a representative Youtube video review to capture public sentiment, comments and feedback.

The choice to rely on Youtube platform was grounded on several pieces of evidence, which show that YouTube serves as a powerful tool due to its massive, diverse user base and the significant role it plays in shaping public discourse and influencing trends.

For instance, in 2022, it was recognized as the most popular social media platform for

[5] https://en.wikipedia.org/wiki/Social_proof

[6] https://opencritic.com

[7] https://www.metacritic.com

gamers in the United States, in regard to gaming[8] (Figure 1).

But it goes beyond just gamers, YouTube reached 2.5 billion monthly active users globally, making it the second most popular website worldwide in 2023 (Figure 2), reaching nearly one-third of all internet users[9].This extensive user base covers numerous countries and cultures, allowing us to capture a broad spectrum of audience reactions to gaming content, and various perspectives on social issues. Additionally, YouTube's audience is well distributed across age groups. Based on data from April 2023, advertisers on YouTube can reach 379.7 million users aged 18 to 24, 522.5 million users aged 25 to 34, and 422 million users aged 35 to 44.[10] This range highlights the platform is used by both younger and middle-aged demographics, perfect for observing the sentiments of both dedicated and casual gamers of every age. By analyzing this varied audience, we can achieve insights that are more representative of the global gaming landscape and understand how diverse groups react to social issues in gaming. In our research we focused on the comment sections of YouTube videos, because they usually enable users to share their views, fostering discussions that can quickly gain virality among large audiences. In the gaming domain the comment sections of videos have become crucial spaces for users to voice their opinions on new videogames and their different aspects, creating a rich source of user-generated content.

Another aspect that makes YouTube an optimal choice for this research is its substantial advertising reach, which in April 2023 was estimated to cover 2.527 billion users globally. [11] Many companies recognize YouTube's potential for reaching diverse audiences and often use the platform to launch trailers, like Ubisoft does, and other promotional content. That is why- as you can read later on the report- we recommend analyzing reactions to such official content as well, where we can gather insights into how different marketing approaches and portrayals of social issues resonate with audiences. For instance, it is four times more likely that viewers turn to YouTube over other platforms to learn about a brand or product, indicating YouTube's influence in shaping consumer perception and brand engagement.[12]

Additionally, YouTube's robust API offers structured data on video attributes such as publish time, video length, category, and regional data. This data accessibility makes YouTube an ideal platform for conducting sentiment analysis and observing audience reactions.

As said before, we focused on analyzing review videos rather than official trailers for several key reasons. Review videos, as user-generated content (UGC), provide more authentic consumer perspectives compared to brand-generated content like official trailers. While trailers are produced by game companies for marketing

[8] statista.com
[9] datareportal.com
[10] datareportal.com
[11] datareportal.com
[12] youtube.com

purposes, review videos are created by actual gamers and consumers, reflecting genuine opinions and experiences with the game. For these reasons we selected the youtubers not only based on their popularity, but also on their authenticity, such as IGN and GameRanx.

Moreover, our choice was underpinned by the fact that review YouTubers often encourage viewers to share their opinions explicitly, inviting them to express their thoughts and experiences. This turns comment sections into an open discussion environment in the comments section. This approach not only enhances community engagement but also establishes a platform where genuine consumer perspectives can thrive. This aligns better with our study's goal of understanding real consumer perceptions and experiences.

Furthermore, A study by Smith, Fischer, and Yongjian (2012) on brand-related UGC across different social media platforms found that UGC tends to be more varied in content and sentiment compared to brand-generated content. This variety in UGC allows us to capture a more diverse range of consumer perspectives and experiences, further justifying our choice of review videos for analysis.

Our thorough choice of videos also involves only considering youtube videos posted shortly after (1 month maximum) the release date of the video game being reviewed. This choice aligns with our purpose of detecting early signs of issues, before they go viral.

For the very same reason, we ran our tests only on comments posted from the day the video was uploaded up to one week afterward (Figure 3).

We started conducting our experimental analysis from 20 datasets containing each video's comments posted in the first week after its upload on Youtube. We first processed comments, using custom stopwords, in order to filter out irrelevant discussions.
Then we conducted Sentiment analysis, and categorized each comment into positive, negative and neutral, with the purpose of obtaining the overall sentiment of the video, that most of the time aligns with the scores of Opencritic and Metacritic.

This analysis was conducted on python using VADER model. This sentiment analysis works by analyzing text to determine whether the sentiment is positive, negative, or neutral using a predefined lexicon, where each word is assigned a sentiment score based on how positive or negative it is. VADER also considers punctuation, capitalization, and the use of intensifiers (like "very") to adjust these scores, recognizing that "I love this!" carries a stronger positive sentiment than "I love this."

In your case, the sentiment is determined based on the polarity score, which is a numerical value that VADER assigns to each text. Depending on this score, VADER categorizes the sentiment as positive, negative, or neutral. This method helps to determine the overall sentiment of comments based on how strongly the text leans towards positive or negative emotion. After this analysis we conducted Topic modeling, in order to obtain the number of topics discussed in each comment section.

We have used the Latent Dirichlet Allocation (LDA) model, which is a probabilistic model that helps uncover hidden thematic structures within a collection of text data. The LDA algorithm works by iteratively adjusting the topic assignments for each word in each document based on the words it co-occurs with, eventually converging on a set of topics that best describe the corpus.

To feed LDA with useful data, you represented the text using TF-IDF (Term Frequency-Inverse Document Frequency), a technique that transforms the raw text into a numerical matrix, where each word is given a weight based on its importance in a specific document relative to its frequency across all documents in the corpus. The intuition behind TF-IDF is that words that appear frequently in a document but rarely across all documents are likely to be more significant. For instance, if a word like "game" appears often in a comment about video games but not in other comments, it will have a higher weight in that comment. This ensures that common but unimportant words (such as "the," "a," or "and") do not dominate the analysis.

A key step we took was the creation of bigrams, which are two-word combinations that frequently appear together. In many cases, single words can be too broad or ambiguous and by identifying common pairs of words we have captured more meaningful concepts or phrases (Figure 4). Finally, we captured the number of topics discussed in each video, using the ‘elbow method’ with a focus on the coherence score. For each number of topics, we computed this coherence score, which is a metric that reflects the interpretability of the topics.

We then plotted the number of topics on the X-axis and the coherence score on the Y-axis and where the graph showed a steady increase in the coherence score up to a certain point and then a noticeable leveling off afterward, we identified the point where the coherence score stopped improving significantly. This point—the elbow—is the most appropriate number of topics to use for the analysis (Figure 5).

All the number of topics were then stored in 2 different datasets: one for well received videogames and one for poorly received. With this arrangement we compared the two datasets using the Mann-Whitney U test, also known as the Wilcoxon rank-sum test, a non-parametric statistical test used to determine whether there is a significant difference between two independent groups.

We decided to use this test because, not only our datasets were independent, but we were also dealing with a numerical measure (number of topics), not normally distributed, as shown by the Shapiro test.
We tested our Null hypothesis **H0**: There is no significant difference in the number of topics between poorly received video games and well-received video games and got a p-value significantly lower than 0.05 (Figure 6). Therefore, we rejected the null hypothesis and concluded that there is a statistically significant difference between the two groups in terms of the number of topics.
This proved our first hypothesis.

To test our second hypothesis, that is: a YouTuber's review sentiment significantly shapes comment sentiment, we employed a rigorous statistical methodology. While it might seem intuitive that the opinion in the comments align with a YouTuber's opinion, our goal was to establish this relationship as statistically significant and measurable, addressing potential concerns about the influence of famous personalities on public reactions.

We used the same 20 datasets containing each video's comment sentiments, used for the previous analysis. In addition to building our dataset, we also went through every single video review and manually extracted the youtuber sentiment. We ended up with a dataset composed of 20 rows and 4 columns: video title, sentiment of comments, youtuber sentiment, and notes (sentences said by the YouTuber in the videos that let the reader understand the sentiment better).

In order to perform statistical tests, the variables of interest were encoded into numerical values: positive sentiments were assigned a value of 1, neutral sentiments a value of 0, and negative sentiments a value of -1. Therefore, we began by exploring the correlation between the YouTuber's sentiment and the aggregated sentiment of the comments. Using Pearson's correlation coefficient, we found a strong and significant positive correlation, with a coefficient of 0.95 and a p-value of less than 0.001 (Figure 7). This demonstrated a clear relationship between the sentiment expressed by the YouTuber and the sentiment in the comment section. However, we know correlation is not causation, so we extended our analysis further by applying regression models.

The first model was a simple linear regression to test whether the YouTuber's sentiment could predict the sentiment in the comments. The regression results were really strong (Figure 8). The coefficient was 1.006, meaning that a one-unit increase in the YouTuber's sentiment (for example, from neutral to positive) led to a nearly equivalent increase in the average sentiment of the comments. Most importantly, the R-squared value was 0.91, indicating that 91% of the variance in comment sentiment could be explained by the YouTuber's sentiment. These findings provided strong evidence in support of our hypothesis that YouTuber sentiment significantly influences comment sentiment.

To be sure about these results we adopted a dual approach running a multinomial logistic regression as well. Our choice is based on the fact that this kind of method accounts for the categorical nature of sentiment, in fact, it is specifically designed for scenarios where the dependent variable consists of more than two categories that do not have a natural order. This model calculates the probability of each category occurring, given the independent variable (in this case, the YouTuber's sentiment). Instead of predicting a single numerical outcome like linear regression, multinomial logistic regression predicts the likelihood of each possible category. For instance, it determines the probability that a comment will be classified as positive, neutral, or negative based on the sentiment expressed by the YouTuber.

From a statistical perspective, this approach allowed us to model the odds of each sentiment category appearing in the comments, which is slightly better for interpretation of the results. The results showed that the model could predict comment sentiment with 95% accuracy, reinforcing the strong relationship between YouTuber sentiment and comment sentiment (Figure 9). Furthermore, the regression coefficients revealed that a positive sentiment from the YouTuber significantly increased the likelihood of positive sentiment appearing in the comments, compared to neutral or negative sentiment.

The convergence and significance of these results suggest that they are logical and not counterintuitive, and similar patterns would likely emerge with larger datasets. The observed relationship between YouTuber sentiment and comment sentiment aligns with expectations, further validating the strength and reliability of our findings.

4. RESULTS:

Our analysis yielded results that provide valuable insights into the relationship between video game reception, YouTube commentary, and audience sentiment.

The first part of our analysis, aimed to test whether there is a significant difference in the number of topics discussed in comments between well-received and poorly received games, yielded clear findings. Using the Mann-Whitney U test, we found a statistically significant difference in the number of topics discussed between the two groups. Specifically, the poorly received video games had a lower variety of topics discussed in their comment sections, suggesting that negative reactions tend to generate narrower and more focused conversations.

In the second part of the analysis, we explored the influence of YouTubers' sentiment on their viewers. We first obtained a strong significant correlation between the sentiment expressed by YouTubers and the sentiment of viewers in the comment sections. This was further tested using a simple linear regression and logistic regression, which confirmed a significant causal relationship between the two variables, supporting our hypothesis that YouTubers’ sentiments do influence the sentiments of their audience.

The robust study aimed at generating significant evidence that underpins undeniably our solution offered to Ubisoft, and from which we can derive valuable recommendations

5. DATA-DRIVEN RECCOMENDATIONS

Based on the previous hypothesis that we verified, Ubisoft should establish a routine, a data-driven process to monitor topic diversity in comments under review videos on platforms like YouTube, particularly within the first week following a game’s launch. As previously stated, a limited range of topics in these discussions can signal early concerns or dissatisfaction.

An effective approach to do so is by setting up automated scripts to scrape and collect comments from review videos of newly released games, focusing on those published by influential YouTubers.

The script should timestamp comments, categorize them by sentiment (positive, neutral, negative), and segment them based on the topics discussed.

In this way, games flagged with low topic diversity could be detected and communicated to relevant teams as soon as possible, enabling proactive responses or quick adjustments to improve player satisfaction and mitigate potential uproar.

Moreover, the strategy could become more robust if implemented across different platforms.

Ubisoft could capture a more comprehensive view of gamer feedback, by extending the sentiment and topic diversity monitoring across multiple platforms, such as Reddit, Twitch, and Twitter.

Our analysis could easily be applied to these channels, but distinct user demographics should be clearly defined, to ensure a meaningful analysis.

For example, Reddit hosts extensive discussion threads on specific games in communities (subreddits), where users can engage in-depth over longer periods. These subreddits allow for deeper, multi-level conversations, which could be more insightful, and rich in terms of user feedback.

Moreover, different subreddits have different audiences, from niche communities to general gaming forums. Therefore it is critical to understand the subreddit's specific demographic and purpose, as this influences discussion tone, topic relevance, and the depth of feedback.

Because of this structure sentiment analysis and topic modeling on Reddit require accounting for both the main comment and the replies, to capture the full context of discussions.

A final thing to consider is Reddit voting system. A system that amplifies popular comments, upvotes, and downvotes may skew visible sentiment, as highly upvoted comments could disproportionately reflect the subreddit's overall mood.

Also, Twitch is a platform that could be highly insightful and rich in uproar signals, because live streaming creates real-time interaction between the streamer and viewers, allowing immediate feedback on games. Along with the fact that Twitch audience is mainly composed of highly involved gamers, which allow to isolate the gamers' opinion, which are the

Our sentiment analysis and topic modeling could be extended to the chat logs and comments made during live-streamed game reviews, to analyze viewer reactions in real-time.

This would imply capturing and segmenting data in shorter intervals to account for changing viewer sentiment throughout the stream. In addition to this, being an environment that fosters free, real-time discussions, it is important to consider the presence of irrelevant comments, like off-topic or meme-based interactions, that should be filtered out to obtain a meaningful sentiment.

Moreover, similar to Youtube, on this platform the streamer's opinion influences viewer sentiment, so it might be hard to isolate organic audience sentiment.

Platform X (formerly Twitter) is valuable as well, specifically for real-time sentiment tracking, especially around game launches or updates. By tracking relevant hashtags, mentions, and keywords associated with a specific game, Ubisoft can quickly identify emerging gamer reactions and observe how specific themes resonate with the wider community.

However, adapting the analysis to X requires a few platform-specific considerations: The platform's character limit encourages concise expressions, which can reduce the diversity of topics compared to platforms that allow longer posts. Therefore, sentiment and topic analysis models should be adapted to effectively interpret short-form expressions, including abbreviations, hashtags, and emojis, to capture a nuanced view of gamer sentiment.

Hashtags play a central role on X, shaping discussions and helping to highlight trending topics. Monitoring shifts in hashtag usage and engagement patterns allows for accurate tracking of theme shifts and trends in gamer sentiment, especially in response to updates or critical moments in a game's lifecycle.

Additionally, X's viral nature can lead to sudden spikes in sentiment around specific issues. By analyzing the patterns of these surges, Ubisoft can differentiate between isolated complaints and broader, sustained trends, helping the response team focus on widespread concerns rather than one-off reactions. Integrating X into a cross-platform sentiment and topic diversity monitoring strategy would provide Ubisoft with timely insights into gamer reactions, complementing data from platforms like Reddit and Twitch to offer a comprehensive view of player feedback across diverse online spaces.

After detecting early uproar signals, we believe that Ubisoft should also implement a proactive response strategy. While we know that the ultimate goal is to prevent problematic events, the data points generated after a game's launch remain valuable, and can help Ubisoft address gamer concerns, maintain a positive brand image, and foster long-term player loyalty.

We know that Ubisoft has already a strong team designated to customer support in place to monitor feedback and engage with players.

The key to success is having prepared a systematic and dedicated approach for these types of situations.

Once the team detects specific issues that players are negatively discussing about, they should issue timely updates via Ubisoft's official social channels and forums to reassure players that their feedback is valued, and that Ubisoft is actively working on solutions.

Along with that, Ubisoft can also prioritize in-game updates and improvements, when possible. This could involve releasing hotfixes for minor issues, rolling out patches for more significant adjustments, or even implementing gameplay changes based on player feedback. We realize that it is not always feasible, but if the signals are detected early enough, this approach could demonstrate Ubisoft's commitment to listening to its community.

We have demonstrated that our analysis of topic modeling is an effective and versatile method to detect early signs of uproar. However, its use is not just restricted to this, topic modeling applied on historical data could in fact identify sensitive themes that have historically sparked negative reactions, such as historical or cultural sensitivities in games. This would allow Ubisoft to either avoid or carefully approach these themes in future games to prevent unintended backlash.

To effectively implement this strategy a list of sensitive themes based on prior game feedback should be drafted and constantly updated with new findings from each game release. This database could guide creative teams during game design, especially when exploring storylines or settings that may be sensitive for certain audiences.

To shift the focus to our second analysis, we have made recommendations also in the strategic marketing domain.

As reported in previous sections, from our analysis, 2 key findings emerged:

There is a high and statistically significant correlation between a YouTuber's opinion and the opinions of the users in the comments.

YouTuber sentiment is a fundamental predictor of the sentiment in the comments.

Based on these results, Ubisoft should go beyond the monitoring of youtubers' opinions, instead it could actively shape the overall sentiment in its favor, by establishing strategic partnerships.

The use of partnerships with influential personalities like influencers, youtubers, actors and famous people in general, is nowadays a widely used practice. However, the criteria to select the correct "celebrity" is not driven just by the number of subscriptions to the channel and views, instead modern research in influencer marketing highlights the importance of matching the brand with the influencer's audience and values to create authentic and impactful campaigns. This is an important aspect of our recommendation in order for Ubisoft to not fall in the so-called Vampire effect, risking that the YouTuber may overshadow the brand and the game thus impairing brand recall and shifting the focus towards himself in the opinions to gather in the comments.

In addition to this, Ubisoft should also consider the youtubers' point of view: many of them are highly viewed for their honesty in reviews, and their credibility often depends on their ability to provide unbiased opinions. To approach this problem, we have made two recommendations.

For post-lunch partnerships, it is crucial to choose the right youtuber based on their preferences, interests and usual video formats, which can help shape the narrative in a way that emphasizes the game's appeal while maintaining influencer authenticity.

Another strategy could be involving early partnerships, before the game's launch. In this phase influencers are usually more likely to accept paid collaborations because it does not compromise their credibility, as they are not yet offering a review but rather previewing or showcasing exclusive content. Ubisoft, in turn, will benefit from this by collecting valuable insights and data from audience feedback during this time window.

These insights can inform product development and marketing strategies, helping to address potential issues before the game's release. That's because the partnership should be implemented early before lunch.

In relation to the pre-release phase, Ubisoft can extend content monitoring analysis in this stage as well. We have noticed that Ubisoft's official trailers of videogames are posted on youtube, making it an ideal video format to apply topic modeling to identify signals of gamer uproar before the game's launch. This approach would give the company time to implement updates or launch targeted marketing campaigns to mitigate potential issues, a significant strategic advantage that can be leveraged.

6. CRITICAL REVIEW: Limitations and Future Research Directions

Our analysis offers robust insights into early signs of gamer uproar and sentiment patterns within the focused scope of 20 video games on YouTube. While the study effectively highlights actionable findings, expanding the dataset size, and incorporating advanced comment filtering techniques would provide deeper and more nuanced insights.
For instance, we are aware that such strong results and significance are already a notable achievement, especially obtained with a relatively small dataset. However, running the analysis on a larger dataset with games that could be strategically representative of different market segments, periods, and varying levels of success and controversy, could make the findings more nuanced.

Furthermore, including a broader array of games would help us explore how different elements, such as game genre, release platform, marketing strategy, and community engagement, might influence public sentiment and topics of discussion. This enriched dataset would provide Ubisoft with a solid empirical foundation to anticipate potential risks and enhance decision-making for future game development and marketing strategies.

Another detail to consider when working with social media data is noise. Irrelevant comments are unrelated to the video game, such as memes and off-topic discussions.
Noise is inevitable, and it is important to handle it with caution, in our case we processed comments removing 'stopwords', to make the final text more relevant to our scope.
A future enhancement we considered is training a classification algorithm with manually labeled comments (relevant-irrelevant). This is a more sophisticated approach, that even if more time-consuming, would have allowed us to have a model capable of accurately distinguishing between noise and relevant comments.

7. CONCLUSION

Our solution empowers Ubisoft to tackle gamer uproar with a proactive, data-driven, and strategic approach that considers the phenomenon's novelty. By leveraging advanced sentiment analysis and topic modeling, we provide an effective early-warning mechanism to detect signals of potential uproar and address it before it escalates. Beyond detection, our framework enables Ubisoft to learn from past controversies by identifying recurring themes, helping the company avoid pitfalls, and designing games that resonate with diverse audiences.

What sets our solution apart is its versatility and actionable focus. From utilizing YouTube as a core platform for gathering insights to recommending strategic influencer partnerships, we provide Ubisoft with tools to shape public sentiment and build stronger connections with both gamers and non-gamers. Expanding this analysis across other platforms like Reddit, Twitch, and X further enhances its robustness, ensuring a comprehensive view of audience reactions.

We are confident that our solution could position Ubisoft to safeguard its reputation, foster trust, and maintain its leadership in the gaming industry. This is not just about addressing uproars, it is about transforming them into opportunities for growth, innovation, and stronger player engagement.

8. APPENDIX

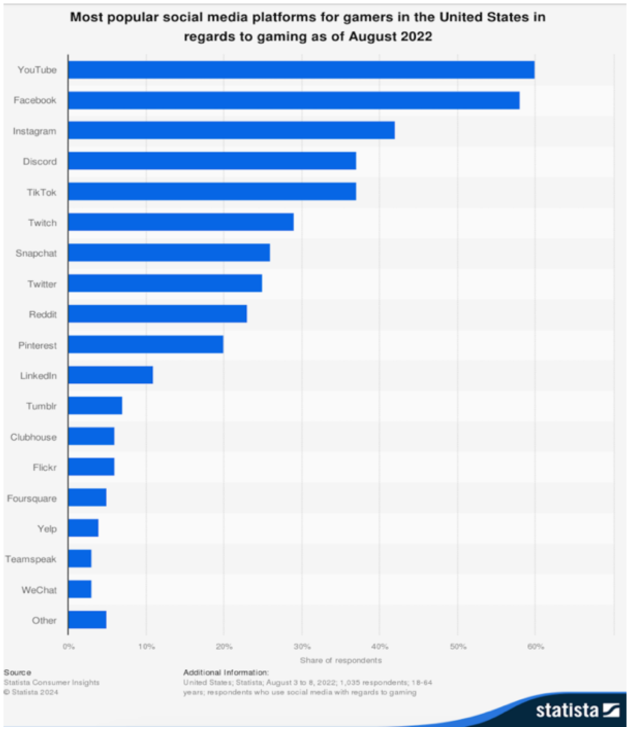

Figure 1: Most popular social media platforms for gamers in the United States in regards to gaming as of August 2022.

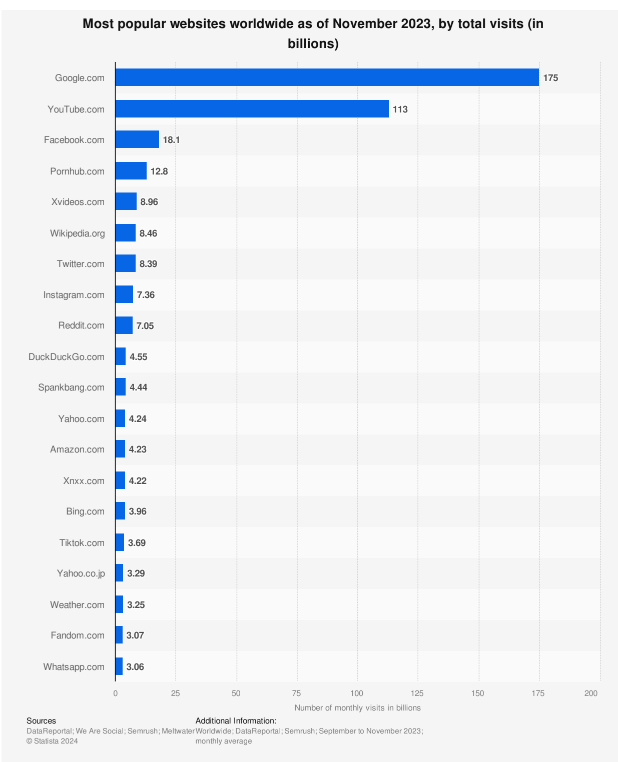

Figure 2: Most popular website worldwide as of November 2023, by total visits (in billions)

title	releaseDate	author	comment	publishedTimeText	replyCount	voteCount
Atomic Heart is 2023's First Major Disappointm...	2023-02-20T13:07:32Z	@underthemayo	Correction: They updated the game to allow dif...	2023-02-20T11:35:02Z	12	104
Atomic Heart is 2023's First Major Disappointm...	2023-02-20T13:07:32Z	@Poopyturdy	Game is so trash holy cow	2024-07-12T03:02:30Z	2	1
Atomic Heart is 2023's First Major Disappointm...	2023-02-20T13:07:32Z	@gabmirdev	LMAO MALUMA	2024-06-21T15:07:54Z	0	0
Atomic Heart is 2023's First Major Disappointm...	2023-02-20T13:07:32Z	@sebastiannunez9023	Under the mayo knows maluma, never thought	2024-06-14T14:57:01Z	0	1

Figure 3: Example of a dataset

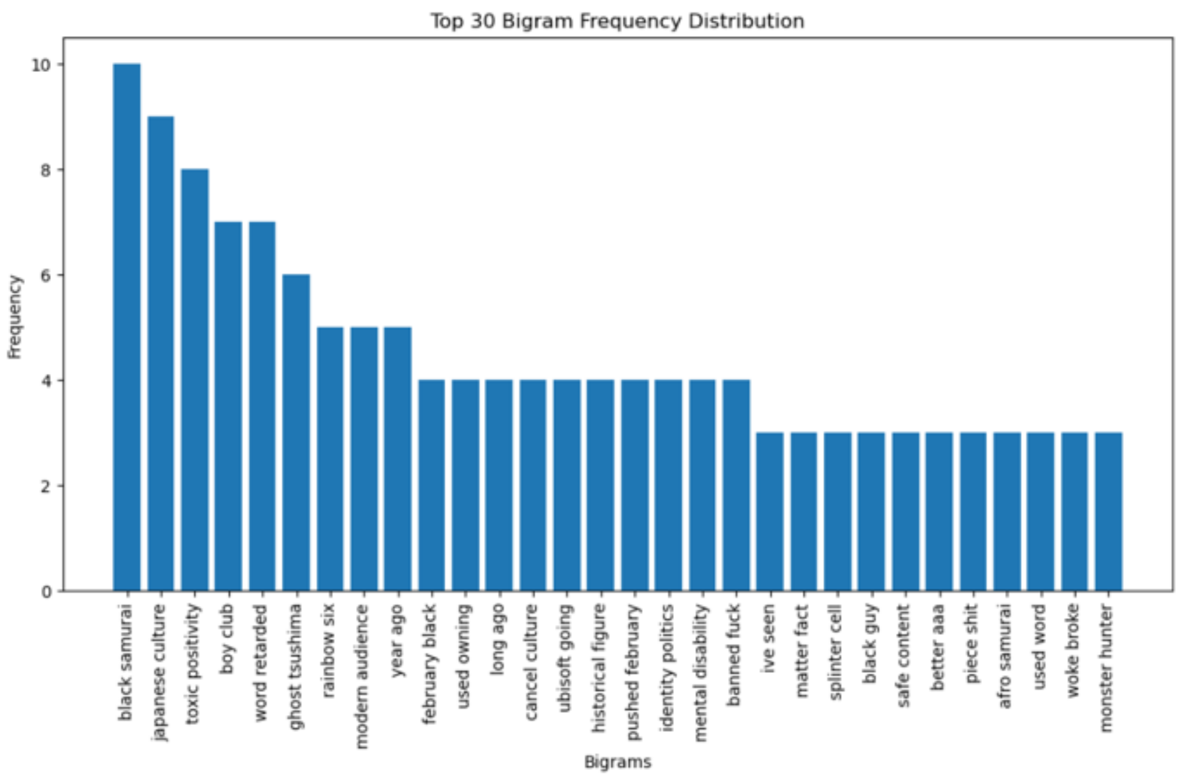

Figure 4: Example of bigrams

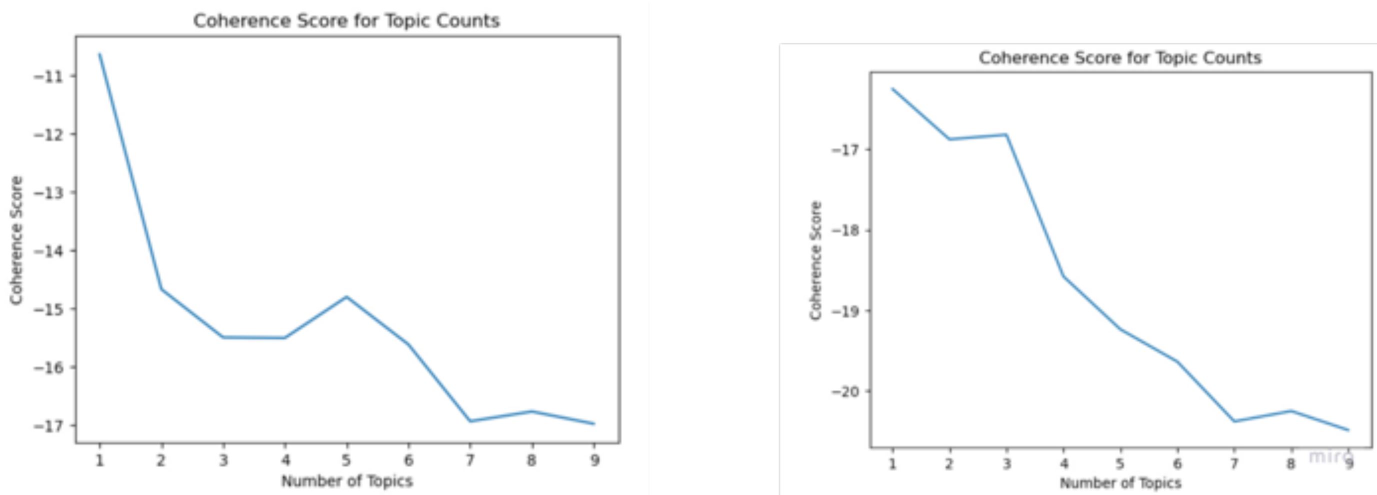

Figure 5: Examples of topic-coherence score graphs

```
Mann-Whitney U Test
Statistic: 93.0
p-value: 0.0009375682445797334
There is a statistically significant difference between the number of topics in good and bad videogames.
```

Figure 6: Mann-Whitney U Test

```
Pearson correlation: 0.9513381762654806
P-value: 1.2172399897818896e-10
```

Figure 7: Correlation between youtuber sentiment and comment section sentiment

```
Linear Regression Coefficient: 1.0058997050147493
Linear Regression Intercept: 0.10029498525073748
R-squared: 0.9050443256201306
```

Figure 8: Linear regression

```
Multinomial Logistic Regression Coefficients: [[-1.47435582]
 [ 0.10821212]
 [ 1.3661437 ]]
Multinomial Logistic Regression Intercept: [ 1.72467683 -1.1846857  -0.53999113]
Multinomial Logistic Regression Accuracy: 0.95
```

Figure 9: Logistic regression